Forecast Bias Correction: A Second-Order Method

Sean Crowell
Department of Mathematics
University of Oklahoma
Norman, OK 73019
scrowell@ou.edu

S. Lakshmivarahan*
School of Computer Science
University of Oklahoma
Norman, OK 73019
varahan@ou.edu

Abstract

In this paper we describe a methodology for determining corrections to parameters and initial conditions in order to improve model forecasts in the presence of data. This methodology is grounded in the variational problem of minimizing the norm of the errors between forecast and data. The general method is presented, and then the method is applied to the specific example of a scalar model, the logistic equation, where it is shown that the method produces satisfactory results.

November 9, 2010

*Corresponding Author: varahan@ou.edu

1 Introduction

Models are abstractions of reality and the goodness of a model is often judged by its ability to explain the observations which are the reflections of the underlying reality. The difference between the actual observation and the model counterpart of the observation is known as the *prediction error* or *forecast bias*. This error/bias is a result of either an "inadequate" model or a result of misspecifications of initial/boundary conditions and/or parameters in the model. Recently, Lakshmivarahan and Lewis (2008) (LL(2008), hereafter) have developed a framework for estimating the error in initial/boundary conditions and/or parameters that will account for the observed forecast bias. Using the (first-order) sensitivity of the solution with respect to the initial conditions and parameters and the observed forecast errors, this framework recasts the bias estimation problem as an inverse problem, in particular as a linear least-squares problem. In LL(2008), they demonstrated the power of this idea using a simplified model consisting of a system of three coupled nonlinear ordinary differential equations that describes the evolution of the mixed-layer over the Gulf of Mexico and actual observations obtained from an earlier field experiment. For details, refer to LL(2008) and references therein.

In this paper we further analyze the power of this framework using a scalar model for population growth known as the logistic model in continuous time. But instead of the first-order method used in LL(2008), in this paper we compare the performance of first-order and second-order methods. For simplicity it is assumed that the model is adequate and the forecast errors are due only to errors in the initial condition and the parameter. As will become evident, the first-order method leads to a linear

1

least-squares problem and the second-order method leads to a nonlinear least-squares problem. Analysis of model inadequacies will be explored in a companion paper.

Section 2 contains a summary of the framework and is adapted from LL(2008). Properties of the logistic model are described in Section 3. Numerical experiments related to the estimation of errors using the first-order and second-order methods along with a comparison of these methods are contained in Section 4. Concluding observations are contained in Section 5.

2 The Framework - a Summary

Let $x(t) \in \mathbb{R}$ denote the state of a dynamical system at time $t \in \mathbb{R}^+$ where $x(0)$ is called the initial state. Let $f : \mathbb{R}\mathsf{x}\mathbb{R}\mathsf{x}\mathbb{R}^+ \to \mathbb{R}$ and the state evolve according to a scalar nonlinear ordinary differential equation given by

$$\dot{x} = f(x, \alpha, t) \tag{2.1}$$

where $\alpha \in \mathbb{R}$ is called the parameter. It is tacitly assumed that the function f in (2.1) satisfies all of the conditions required for the existence, uniqueness, and smoothness of the solution $x(t) = x(t, \alpha, x(0))$ for all $t \in \mathbb{R}^+$.

Let $h : \mathbb{R} \to \mathbb{R}$ be a smooth function that defines the model counterpart of the observation

$$z_t^M = h(x(t)) \tag{2.2}$$

Let z_t be the actual observation obtained from the field measurement. Then

$$e_t = z_t - z_t^M \tag{2.3}$$

denotes the prediction error or the forecast bias. It is assumed that the model in (2.1) does not have any deficiencies and the forecast bias is mainly due to the misspecification of the initial condition, $x(0)$, and/or the parameter α.

Let Δx be the actual change in $x(t)$ induced by the perturbations $\delta x(0)$ in $x(0)$ and $\delta \alpha$ in α. The goal is to find $\beta = (\delta x(0), \delta \alpha)^{\mathsf{T}}$, the vector of perturbations such that

$$z_t - h(x_t + \Delta x) = 0. \tag{2.4}$$

Let Δh be the actual change in $h(x(t))$ induced by the change Δx in $x(t)$. Then

$$h(x(t) + \Delta x) = h(x(t)) + \Delta h \tag{2.5}$$

Combining (2.3) - (2.5), our goal is to find β such that

$$e_t = \Delta h \tag{2.6}$$

Recall that

$$\Delta h = \sum_{j=1}^{\infty} \delta^j h \tag{2.7}$$

where $\delta^j h$ is called the $j^t h$ variation of h, the fraction of the induced change that is

3

attributable to the jth derivative of h and the change in Δx in $x(t)$. Similarly

$$\Delta x = \sum_{j=1}^{\infty} \delta^j x \tag{2.8}$$

where $\delta^j x$, called the jth variation of x, is the fraction of the total change in $x(t)$ that is attributable to the jth partial derivatives of $x(t)$ with respect to α and $x(0)$ and β.

In practice we approximate the infinite sum by taking only the first k terms, resulting in a kth-order approximation given by

$$\Delta h = \sum_{j=1}^{k} \delta^j h \text{ and } \Delta x = \sum_{j=1}^{k} \delta^j x. \tag{2.9}$$

2.1 First-order Approximation

Setting $k = 1$, we get

$$\Delta h = \delta h \text{ and } \Delta x = \delta x. \tag{2.10}$$

Let $D_a(g) = \frac{\partial g}{\partial a}$. From first principles and (2.10), it follows that

$$\delta h = D_x(h)\Delta x = D_x(h)\delta x \tag{2.11}$$

and

$$\delta x = D_{x(0)}(x)\delta x(0) + D_\alpha(x)\delta\alpha. \tag{2.12}$$

4

The first derivative $D_a(g)$ is called the first-order sensitivity of g with respect to a (Cacuci (2003) and Cruz(1973)). Setting

$$\overline{H_1} = D_{x(0)}(x), \ \overline{H_2} = D_\alpha(x), \overline{H} = [\overline{H_1}, \overline{H_2}] \in \mathbb{R}^{1\times 2},$$

(2.12) can be rewritten as

$$\delta x = \overline{H}\beta. \tag{2.13}$$

Combining (2.13) and (2.11) with (2.6) we obtain an underdetermined linear least-squares problem

$$H\beta = e \tag{2.14}$$

where

$$H = D_x(h)\overline{H} \in \mathbb{R}^{1\times 2} \tag{2.15}$$

Suppose that there are N observations $z_{t_1}, z_{t_2}, \ldots z_{t_N}$ available at times $0 \le t_1 < t_2 < \ldots < t_N$. Then, at each time t_i, $1 \le i \le N$, we have the forecast error

$$e_{t_i} = z_{t_i} - h(x_{t_i}) \tag{2.16}$$

Define

$$\overline{H_1}(t_i) = D_{x(0)}(x(t_i)), \ \overline{H_2}(t_i) = D_\alpha(x(t_i)) \tag{2.17}$$
$$\overline{H}(t_i) = [\overline{H_1}(t_i), \overline{H_2}(t_i)] \in \mathbb{R}^{1\times 2}$$

5

and a diagonal matrix

$$D(h) = \text{Diag}\{D_1(h), D_2(h), ..., D_N(h)\} \in \mathbb{R}^{N \times N}$$

where

$$D_i(h) = D_x(h(x(t_i))).$$

for simplicity in notation. Let

$$\overline{H_N} = \begin{bmatrix} \overline{H_1}(t_1) & \overline{H_2}(t_1) \\ \overline{H_1}(t_1) & \overline{H_2}(t_1) \\ \vdots & \vdots \\ \overline{H_1}(t_N) & \overline{H_2}(t_N) \end{bmatrix} \in \mathbb{R}^{N \times 2} \tag{2.18}$$

$$H_N = D(h)\overline{H_N} \tag{2.19}$$

and

$$E_N = (e_{t_1}, e_{t_2}, ..., e_{t_N})^\mathsf{T} \in \mathbb{R}^N \tag{2.20}$$

For this case of N observations, in place of (2.14), we obtain a linear least-squares problem

$$H_N\beta = E_N \tag{2.21}$$

where $H_N \in \mathbb{R}^{N \times 2}$, $\beta \in \mathbb{R}^2$, and $E_N \in \mathbb{R}^N$.

The unknown β is obtained by minimizing

$$g_1(\beta) = \|H_N\beta - E_N\|^2 \tag{2.22}$$

6

where it is tacitly assumed that the matrix H_N is of full rank. Under this assumption, β is given by (Lewis, et al (2006))

$$\beta = \begin{cases} \left(H_N^\mathsf{T} H_N\right)^{-1} H_N^\mathsf{T} E_N & \text{if } N > 2 \\ H_N^{-1} E_N & \text{if } N = 2 \\ H_N^\mathsf{T} \left(H_N H_N^\mathsf{T}\right)^{-1} E_N & \text{if } N < 2 \end{cases} \tag{2.23}$$

When H_N is not of full rank, we invoke the Tikhonov regularization (Lewis, et al (2006)) using which β is obtained by minimizing

$$\overline{g_1}(\beta) = \|H_N \beta - E_N\|^2 + \frac{\lambda^2}{2} \|\beta\|^2 \tag{2.24}$$

for some real constant $\lambda > 0$, called the regularization parameter. The minimizing β in this case is given by

$$\overline{\beta}_{LS} = (H_N^\mathsf{T} H_N + \lambda I)^{-1} H_N^\mathsf{T} E_N \tag{2.25}$$

In place of Tikhonov regularization, one could use the generalized inverse of the matrix H_N to obtain

$$\beta = H_N^+ E_N$$

where H_N^+ is the Moore-Penrose inverse (Lewis et al (2006)).

2.2 Second-Order Approximation

Setting $k = 2$, in (2.9), it follows that

$$\Delta x = \delta x + \delta^2 x \tag{2.26}$$

where δx is given by (2.13) is linear in β. From first principles, we obtain

$$\delta^2 x = \frac{1}{2} \beta^{\mathsf{T}} D^2(x) \beta \tag{2.27}$$

which is quadratic in β and $D^2(x)$ is the Hessian of $x(t)$ with respect to the entries of β, that is

$$D^2(x) = \begin{bmatrix} D^2_{x(0)}(x(t)) & D^2_{x(0),\alpha}(x(t)) \\ D^2_{\alpha,x(0)}(x(t)) & D^2_{\alpha}(x(t)) \end{bmatrix} \tag{2.28}$$

where $D^2_a(g) = \frac{\partial^2 g}{\partial a^2}$ and $D^2_{a,b}(g) = \frac{\partial^2 g}{\partial a \partial b}$. The second partial derivatives of g are also called the second-order sensitivities of g with respect to a and b. Similarly, from (2.9),

$$\Delta h = \delta h + \delta^2 h = D_x(h)\Delta x + \frac{1}{2} D^2_x(h)(\Delta x)^2 \tag{2.29}$$

Substituting (2.26) in (2.29) and using (2.6) for the case of the single observation, we obtain a nonlinear least-squares problem.

$$e_t = D_x(h)(\delta x + \delta^2 x) + \frac{1}{2} D^2_x(h)(\delta x + \delta^2 x)^2 \tag{2.30}$$

The unknown β is then obtained by minimizing

$$g_2(\beta) = \left\| e_t - D_x(h)(\delta x + \delta^2 x) + \frac{1}{2}D_x^2(h)(\delta x + \delta^2 x)^2 \right\|^2 \tag{2.31}$$

In the special case where the state is directly observable, $h(x) = x$ and so $D_x(h) = 1$ and $D_x^2(x) = 0$. Substituting these into (2.31), in view of (2.13) and (2.17), we get

$$g_2(\beta) = \left\| e_t - \overline{H}\beta - \frac{1}{2}\beta^\mathsf{T} D^2(x)\beta \right\|^2 \tag{2.32}$$

Clearly, $g_2(\beta)$ is a fourth degree polynomial in the components of β and in general can have multiple minima which could further complicate the minimization problem. When there are N (≥ 2) observations at times $0 \leq t_0 < t_1 < \cdots < t_{N-1}$, then we get N versions of the relation (2.30), one for each time. The unknown β is then obtained by minimizing

$$G_2(\beta) = \sum_{i=0}^{N-1} \left\| e_{t_i} - D_{x(t_i)}(h)[\delta x(t_i) + \delta^2 x(t_i)] - \frac{1}{2}D_{x(t_i)}^2(h)[\delta x(t_i) + \delta^2 x(t_i)]^2 \right\|^2. \tag{2.33}$$

In the special case when $h(x) = x$, $D_x(h) = 1$ and $D_x^2(x) = 0$ and (2.33) reduces to

$$G_2(\beta) = \sum_{i=0}^{N-1} \left\| e_{t_i} - \overline{H}(t_i)\beta - \frac{1}{2}\beta^\mathsf{T} D^2(x(t_i)\beta \right\|^2. \tag{2.34}$$

9

3   The Logistic Equation

Consider the standard logistic equation that describes the growth of a certain population in an environment with carrying capacity 1 and growth rate parameter $\alpha > 0$ given by

$$\dot{x} = \alpha x(1 - x) \tag{3.1}$$

with $x(0) = x_o > 0$ as the initial condition. The solution of this nonlinear ordinary differential equation is given by

$$x(t) = \frac{x_o e^{\alpha t}}{1 - x_o + x_o e^{\alpha t}} \tag{3.2}$$

Notice that $x(t)$ depends nonlinearly on x_o and α.

It can be verified that $x(t) \to 1$ as $t \to \infty$ and $x(t) \in [0, 1]$ for all t whenever $x(0) \in [0, 1]$, for $\alpha > 0$. Since $\dot{x}(t) > 0$ for all t, and so is an increasing function which increases from x_o to 1 as t increases from 0 to ∞. A typical plot of $x(t)$ in (3.2) for $x(0) = 0.5$ and $\alpha = 1$ is given in Figure 1.

The various first-order sensitivities of the the solution with respect to $\beta = (x(0), \alpha)^{\mathsf{T}}$ are given by

$$D_{x(0)}(x) = \frac{e^{\alpha t}}{[1 - x_o + x_o e^{\alpha t}]^2} \tag{3.3}$$

$$D_{\alpha}(x) = \frac{x_o(1 - x_o)t e^{\alpha t}}{[1 - x_o + x_o e^{\alpha t}]^2} = x_o(1 - x_o)t D_{x(0)}(x) \tag{3.4}$$

That is, $D_\alpha(x)$ is a multiple of $D_{x(0)}$. Further it can be verified for large t that

$$D_\alpha(x) \approx \frac{1 - x_o}{x_o} t e^{-\alpha t} \text{ and } D_{x(0)}(x) \approx \frac{1}{x_o{}^2} e^{-\alpha t}$$

Thus, $D_\alpha(x)$ and hence $D_{x(0)}(x)$ tend to zero as $t \to \infty$. A plot of $D_\alpha(t)$ and $D_{x(0)}(x)$ versus t for $x_o = 0.5$ and $\alpha = 1$ is given in Figure 2. This in turn implies that $x(t)$ becomes less sensitive with respect to α and $x(0)$ for large t, and that for our framework to be effective, the observations have to be taken during the transient phase of the solution.

The Hessian of $x(t)$ with respect to β whose elements are the second-order sensitivity functions are given by

$$D^2(x) = \begin{bmatrix} D^2_{x(0)}(x) & D^2_{x(0),\alpha}(x) \\ D^2_{\alpha,x(0)}(x) & D^2_\alpha(x) \end{bmatrix} \tag{3.5}$$

where

$$D^2_{x(0)}(x) = \frac{2(e^{\alpha t} - e^{2\alpha t})}{[1 - x_o + x_o e^{\alpha t}]^3}$$

$$D^2_\alpha(x) = \frac{x_o(1 - x_o)t^2 e^{\alpha t}[1 - x_o - x_o e^{\alpha t}]}{[1 - x_o + x_o e^{\alpha t}]^3}$$

and

$$D^2_{x(0),\alpha}(x) = D^2_{\alpha,x(0)}(x) = \frac{t^2 e^{\alpha t}[1 - x_o - x_o e^{\alpha t}]}{[1 - x_o + x_o e^{\alpha t}]^3}.$$

It is assumed that $x(t)$ is directly observable, and so $h(x) = x$. Accordingly $D_x(h) \equiv 1$ and $D^2_x(h) \equiv 0$.

11

3.1 The First-Order Approximation

In this case (2.14) reduces to (2.13) and from (3.3) and (3.4) we get

$$H = \overline{H} = D_{x(0)}(x)[1, x_0(1 - x_0)t] \in \mathbb{R}^{1 \times 2}. \tag{3.6}$$

Two cases arise. The first case is when there are less observations than parameters. The second is when the number of observations is at least as great as the number of unknown parameters. In this paper, we only address the latter case. In this case, $N > 1$ and

$$H_N = \begin{bmatrix} D_{x(0)}(x(t_1)) & D_{x(0)}(x(t_1))x_o(1 - x_o)t_1 \\ D_{x(0)}(x(t_2)) & D_{x(0)}(x(t_2))x_o(1 - x_o)t_2 \\ \vdots & \vdots \\ D_{x(0)}(x(t_N)) & D_{x(0)}(x(t_N))x_o(1 - x_o)t_N \end{bmatrix} \tag{3.7}$$

It can be verified that matrix $H_N \in \mathbb{R}^{N \times 2}$ is of rank two and the least-square solution is given by

$$\beta_{LS}^f = \begin{cases} H_N^{-1}E_N & if \ N = 2 \\ (H_N^{\mathsf{T}}H_N)^{-1}H_N^{\mathsf{T}}E_N & if \ N > 2 \end{cases} \tag{3.8}$$

3.2 The Second-Order Approximation

Expanding the right hand side of (2.34) and dropping the terms of degree 3 and 4, we obtain a quadratic approximation $Q(\beta)$ to $g_2(\beta)$ given by

$$Q(\beta) = \sum_{i=0}^{N-1} \left[e_{t_i}^2 - 2e_{t_i}\overline{H}(t_i)\beta + \beta^{\mathsf{T}}[\overline{H}(t_i)^{\mathsf{T}}\overline{H}(t_i) - e_{t_i}D^2(x(t_i))]\beta \right]. \tag{3.9}$$

12

The β that minimizes this is given by

$$\beta_{LS}^s = \sum_{i=0}^{N-1} \left[\overline{H}(t_i)^{\mathsf{T}} \overline{H}(t_i) - e_{t_i} D^2(x(t_i)) \right] \overline{H}(t_i)^{\mathsf{T}} e_{t_i}. \qquad (3.10)$$

4   Numerical Experiments

In this section we compare the performance of the first-order and the second-order methods using N (≥ 2) observations.

4.1   Generation of Observations

It is assumed that $h(x) = x$, that is, the state of the system $x(t)$ is directly observable and that $z_t = x(t)$. We use the model solution $\overline{x}(t)$ starting from $\overline{x}(0)$ and $\overline{\alpha}$ to generate different sets of N observations.

The distribution of the N observation times

$$t_0 < t_1 < t_2 < \cdots < t_{N-1}$$

are given in Table 1 where $t_i = t_0 + i(k\Delta)$ with $\Delta = 0.5$ (fixed), $k = 1, 4, 8$ and 12 and $0 \leq i \leq N - 1$. Here t_0 denotes the starting time and $(k\Delta)$ denotes the time between successive observations. Thus, for example, for $t_0 = 0$ and $k = 4$, we generate $N = 6$ observations at times $t_0 = 0.0$, $t_1 = 2.0$, $t_2 = 4.0$, $t_3 = 6.0$, $t_4 = 8.0$, and $t_5 = 10.0$, which corresponds to the second row of Table 1.

The actual observations $z_{t_i} = \overline{x}(t_i)$ where $\overline{x}(t_i)$ is the solution of (3.1) starting from $\overline{x}(0) = 0.5$ and $\overline{\alpha} = 1.0$ used in the estimation of β are given in Table 2.

13

4.2 A Plan for the Experiment

Let $x(t)$ be the solution of the model (3.1) starting from $x(0) \neq \overline{x}(0)$ and $\alpha \neq \overline{\alpha}$. Our goal is to estimate $\delta x(0) = x(0) - \overline{x}(0)$ and $\delta \alpha = \alpha - \overline{\alpha}$ using the set of observations in Table 2. We divide the experiment into three parts based on N, the number of observations.

<u>Case 1</u>: $N = 2$. We compute the estimates of β using two observations at times t_0 and $t_0 + k\Delta$ for $k = 1$, 4, 8 and 12 and for $t_0 = 0$, 4 and 8. Estimates of β using the first-order and the second-order methods are given in the top part of Tables 3 and 4 respectively.

For example, referring to the top part of Table 3, the estimate of $\beta = (\delta x(0) \; \delta \alpha)$ using two observations $z_{t_0} = z_4 = x(4)$ and $z_{t_1} = z_{4.5} = x(4.5)$ and the first-order method is given by $\beta = (-0.0955, 0.0981)$. Since the actual value of $\delta x(0) = -0.1$ and $\delta \alpha = 0.1$, the relative error in the estimate of $\delta x(0) = \frac{-0.0955 - (-0.1)}{-0.1} = 0.045$ and that of $\delta \alpha = \frac{0.0981 - 0.1}{0.1} = -0.019$. Similarly for all other patterns of $N = 2$ observations.

<u>Case 2</u>: $N = 4$. Results of the estimation of β using 4 observations at times t_0, $t_0 + k\Delta$, $t_0 + 2k\Delta$ and $t_0 + 3k\Delta$ for $k = 1$, 4, 8 and 12 and $t_0 = 0$, 4 and 8 are given in the middle parts of Tables 3 and 4.

<u>Case 3</u>: $N = 6$. Results of the estimation of β using 4 observations at times t_0, $t_0 + k\Delta$, $t_0 + 2k\Delta$, ..., $t_0 + 5k\Delta$ for $k = 1$, 4, 8 and 12 and $t_0 = 0$, 4 and 8 are given in the bottom parts of Tables 3 and 4.

14

4.3 Comments and Discussion

A number of observations are in order.

1. When two observations are too close to each other, say by a distance $\Delta = 0.5$, the solutions at these two times are very close to each other and consequently the sensitivity functions at time t_0 $(D_{x(0)}(x(t_0)),\, D_\alpha(x(t_0)))$ and at time $t_1 = t_0 + \Delta$ namely $(D_{x(0)}(x(t_1)),\, D_\alpha(x(t_1)))$ are also very close to each other. This in turn implies that the two rows in $\overline{H}_N$ are very nearly colinear, which leads to bad conditioning of the matrix $\overline{H}_N{}^{\mathsf{T}}\overline{H}_N$. Hence there is a greater chance for the estimates to be less reliable.

2. From (3.2) and Figure 1, it follows that the solution $x(t)$ attains the steady state value which is independent of $x(0)$ and α. In other words, the sensitivity of $x(t)$ with respect to $x(0)$ and α decrease to zero as t increases. Thus, from (3.3) and (3.4) it follows that $D_{x(0)}(x(t))$ and $D_\alpha(x(t))$ both tend to zero as t increases. Thus, the row of $\overline{H}_N$ for large t becomes zero leading to ill-conditioning of $\overline{H}_N$ which makes the estimates less reliable.

3. Since the solution $x(t)$ depends on $x(0)$ and α nonlinearly, we cannot hope to recover the actual perturbations δx_0 and $\delta \alpha$ that led to the forecast bias in the first place. While in principle the second-order method is better than the first-order counterpart, to improve the overall accuracy of the estimate, one may have to resort to an iterative improvement of the estimate. A detailed comparison of the performance of the iterative versions of the first and second order methods for small values of the perturbations ($\delta x_0 = -0.1$ and $\delta \alpha = 0.1$)

are given in Tables 5 and 6 respectively. Tables 7 and 8 contain similar results for large values of δx_0 and $\delta \alpha$.

It turns out that the iterative versions of the first-order method peform at least as well as the iterative versions of the second-order counterpart. Given that second-order methods require larger computation, from this exercise it follows that the iterative first-order method would be a good choice for estimating the forecast bias.

5   References

Cacuci, D.G. (2003) *Sensitivity and Uncertainty Analysis*, Chapman and Hall/CRC Press, Boca Raton, FL

Cruz, Jr. J.B. (1973) *System Sensitivity Analysis*, Stroudsberg, PA. (editor)

Lakshmivarahan, S. and J. Lewis (2008) "Finding sources of bias error in forecast models: A framework," *Monthly Weather Review.* (submitted)

Lewis, J., S. Lakshmivarahan, and S.K. Dhall (2006) *Dynamic Data Assimilation*, Cambridge University Press, 654 pages
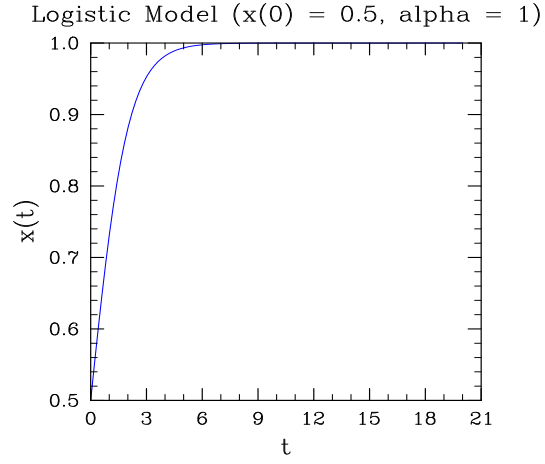
Figure 1: Typical Solution of Logistic Equation ($x(0) = 0.5$ and $\alpha = 1$)

(a) Sensitivity with respect to x(0)

(b) Sensitivity with respect to α
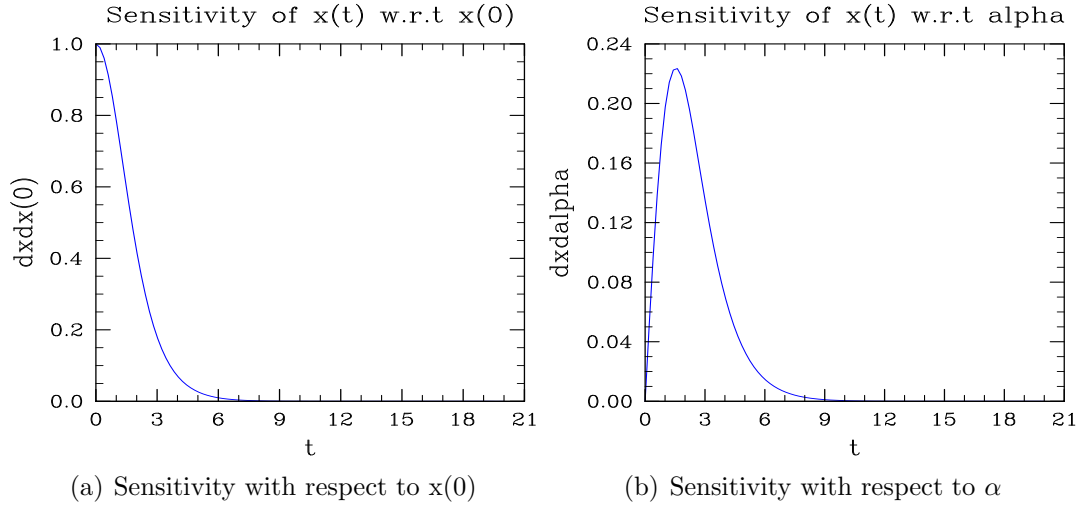
Figure 2: First-Order Sensitivities of the Solution to the Logistic Equation

Observation Times			
$k\Delta$	Δ	4Δ	8Δ	12Δ
t_0	0.0	0.0	0.0	0.0
t_1	0.5	2.0	4.0	6.0
t_2	1.0	4.0	8.0	12.0
t_3	1.5	6.0	12.0	18.0
t_4	2.0	8.0	16.0	24.0
t_5	2.5	10.0	20.0	30.0
$k\Delta$	Δ	4Δ	8Δ	12Δ
t_0	4.0	4.0	4.0	4.0
t_1	4.5	6.0	8.0	10.0
t_2	5.0	8.0	12.0	16.0
t_3	5.5	10.0	16.0	22.0
t_4	6.0	12.0	20.0	28.0
t_5	6.5	14.0	24.0	34.0
$k\Delta$	Δ	4Δ	8Δ	12Δ
t_0	8.0	8.0	8.0	8.0
t_1	8.5	10.0	12.0	14.0
t_2	9.0	12.0	16.0	20.0
t_3	9.5	14.0	20.0	26.0
t_4	10.0	16.0	24.0	32.0
t_5	10.5	18.0	28.0	38.0

Table 1: Distribution of the N observation times: $t_i = t_0 + k(i-1)\Delta$ with $\Delta = 0.5$, $k = 1, 4, 8$, and 12 and $0 \leq i \leq N-1$. t_0 denotes the starting time and $k\Delta$ denotes the time interval between successive observations and N denotes the number of observations.

Actual Observations			
Multiple of Δ			
$t_0 = 0$	1	4	8	12
z_0	0.50000000000000	0.50000000000000	0.50000000000000	0.50000000000000
z_1	0.62245933120185	0.88079707797788	0.98201379003791	0.99752737684337
z_2	0.73105857863000	0.98201379003791	0.99966464986953	0.99999385582540
z_3	0.81757447619364	0.99752737684337	0.99999385582540	0.99999998477002
z_4	0.88079707797788	0.99966464986953	0.99999988746484	0.99999999996225
z_5	0.92414181997876	0.99995460213130	0.99999999793885	0.99999999999991
Multiple of Δ			
$t_0 = 4$	1	4	8	12
z_0	0.98201379003791	0.98201379003791	0.98201379003791	0.98201379003791
z_1	0.98901305736941	0.99752737684337	0.99966464986953	0.99995460213130
z_2	0.99330714907572	0.99966464986953	0.99999385582540	0.99999988746484
z_3	0.99592986228410	0.99995460213130	0.99999988746484	0.99999999972105
z_4	0.99752737684337	0.99999385582540	0.99999999793885	0.99999999999931
z_5	0.99849881774326	0.99999916847197	0.99999999996225	1.00000000000000
Multiple of Δ			
$t_0 = 8$	1	4	8	12
z_0	0.99966464986953	0.99966464986953	0.99966464986953	0.99966464986953
z_1	0.99979657302194	0.99995460213130	0.99999385582540	0.99999916847197
z_2	0.99987660542401	0.99999385582540	0.99999988746484	0.99999999793885
z_3	0.99992515377249	0.99999916847197	0.99999999793885	0.99999999999489
z_4	0.99995460213130	0.99999988746484	0.99999999996225	0.99999999999999
z_5	0.99997246430889	0.99999998477002	0.99999999999931	1.00000000000000

Table 2: Actual Observations

True Corrections		$\delta x_o = -0.1$		$\delta\alpha = 0.1$	
2 observations		multiple of Δ			
t_o		1	4	8	12
0	δx_o	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.0761	0.0968	0.1028	0.0989
	κ	2.8e+2	4.1e+1	2.0e+2	3.1e+3
4	δx_o	-0.0955	-0.0888	-0.0809	-0.0739
	$\delta\alpha$	0.0981	0.0912	0.0829	0.0756
	κ	1.3e+3	6.0e+2	5.0e+3	8.1e+4
8	δx_o	-0.0479	-0.0390	-0.0284	-0.0190
	$\delta\alpha$	0.0657	0.0611	0.0555	0.0507
	κ	7.9e+3	3.8e+3	3.2e+4	5.2e+5
4 observations		multiple of Δ			
t_o		1	4	8	12
0	δx_o	-0.1006	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.0892	0.0977	0.1028	0.0989
	κ	7.0e+1	3.5e+1	2.0e+2	3.1e+3
4	δx_o	-0.1038	-0.1108	-0.1477	0.4009
	$\delta\alpha$	0.0977	0.1039	0.1349	-0.0570
	κ	3.9e+2	9.4e+2	5.0e+4	1.5e+5
8	δx_o	-0.0440	-0.0378	-0.0283	-0.0190
	$\delta\alpha$	0.0637	0.0605	0.0555	0.0507
	κ	2.5e+3	3.4e+3	3.2e+4	5.2e+5
6 observations		multiple of Δ			
t_o		1	4	8	12
0	δx_o	-0.1011	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.0951	0.0977	0.1028	0.0989
	κ	4.2e+1	3.5e+1	2.0e+2	3.1e+3
4	δx_o	-0.0907	-0.0879	-0.0808	-0.0739
	$\delta\alpha$	0.0932	0.0902	0.0828	0.0756
	κ	3.0e+2	5.4e+2	5.0e+3	8.1e+4
8	δx_o	-0.0416	-0.0378	-0.0283	-0.0190
	$\delta\alpha$	0.0625	0.0605	0.0555	0.0507
	κ	1.8e+3	3.4e+3	3.2e+4	5.2e+5

Table 3: Estimates of β using the first-order method. κ denotes the condition number of the matrix $\overline{H}_N^\mathsf{T}\overline{H}_N$, where N is the number of observations used in the estimation.

True Corrections		$\delta x_o = -0.1$		$\delta\alpha = 0.1$	
2 observations		multiple of Δ			
t_o		1	4	8	12
0	δx_o	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1117	0.0987	0.0945	0.0980
	κ	2.1e+2	2.5e+1	1.8e+2	7.1e+3
4	δx_o	-0.1009	-0.1098	-0.1489	-0.1264
	$\delta\alpha$	0.0953	0.1030	0.1359	0.1247
	κ	1.1e+3	9.6e+2	5.1e+4	2.5e+5
8	δx_o	0.3763	0.3744	0.3737	0.3749
	$\delta\alpha$	-0.0528	-0.0523	-0.0521	-0.0524
	κ	8.2e+3	4.7e+3	5.2e+4	1.1e+6
4 observations		multiple of Δ			
t_o		1	4	8	12
0	δx_o	-0.0995	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1025	0.0983	0.0945	0.0980
	κ	5.4e+1	2.3e+1	1.8e+2	7.1e+3
4	δx_o	-0.1038	-0.1108	-0.1477	0.4009
	$\delta\alpha$	0.0977	0.1039	0.1349	-0.0570
	κ	3.9e+2	9.4e+2	5.0e+4	1.5e+5
8	δx_o	0.3754	0.3743	0.3737	0.3749
	$\delta\alpha$	-0.0525	-0.0523	-0.0521	-0.0524
	κ	2.6e+3	4.3e+3	5.2e+4	1.1e+6
6 observations		multiple of Δ			
t_o		1	4	8	12
0	δx_o	-0.0991	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.0991	0.0983	0.0945	0.0980
	κ	3.3e+1	2.3e+1	1.8e+2	7.1e+3
4	δx_o	-0.1059	-0.1108	-0.1477	0.4009
	$\delta\alpha$	0.0995	0.1039	0.1349	-0.0570
	κ	3.0e+2	9.4e+2	5.0e+4	1.5e+5
8	δx_o	0.3750	0.3743	0.3737	0.3749
	$\delta\alpha$	-0.0524	-0.0523	-0.0521	-0.0524
	κ	1.9e+3	4.3e+3	5.2e+4	1.1e+6

Table 4: Estimates of β using the second-order method. κ denotes the condition number of the matrix $\overline{H}_N^{\mathsf{T}}\overline{H}_N$, where N is the number of observations used in the estimation.

True Corrections		$\delta x_o = -0.1$		$\delta\alpha = 0.1$	
2 observations		multiple of Δ			
t_0		1	4	8	12
0	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	2.6e+2	3.2e+1	2.0e+2	4.6e+3
	iterations	4	4	4	4
4	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	1.4e+3	8.7e+2	1.1e+4	2.7e+5
	iterations	4	4	4	5
8	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	8.6e+3	5.7e+3	7.5e+4	1.8e+6
	iterations	5	5	5	5
4 observations		multiple of Δ			
t_o		1	4	8	12
0	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	6.3e+1	2.9e+1	2.0e+2	4.6e+3
	iterations	4	4	4	4
4	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	4.6e+2	8.1e+2	1.1e+4	2.7e+5
	iterations	4	4	4	5
8	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	2.9e+3	5.3e+3	7.4e+4	1.8e+6
	iterations	5	5	5	5
6 observations		multiple of Δ			
t_o		1	4	8	12
0	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	3.9e+1	2.9e+1	2.0e+2	4.6e+3
	iterations	4	4	4	4
	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	3.5e+2	8.1e+2	1.1e+4	2.7e+5
	iterations	4	4	4	5
8	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	2.2e+3	5.3e+3	7.4e+4	1.8e+6
	iterations	5	5	5	5

Table 5: Estimates of β using the iterated first-order method. κ denotes the condition number of the matrix $\overline{H}_N^{\mathsf{T}}\overline{H}_N$, where N is the number of observations used in the estimation. The number of iterations is that required to reach a threshold of 10^{-6} (at most 10)

True Corrections		$\delta x_o = -0.1$		$\delta\alpha = 0.1$	
2 observations		multiple of Δ			
t_0		1	4	8	12
0	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	2.6e+2	3.2e+1	2.0e+2	4.6e+3
	iterations	4	4	4	4
4	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	1.4e+3	8.7e+2	1.1e+4	2.7e+5
	iterations	4	4	5	6
8	δx_0	-0.1000	-0.1000	-0.1000	NaN
	$\delta\alpha$	0.1000	0.1000	0.1000	NaN
	κ	8.6e+3	5.7e+3	7.5e+4	1.0e+9
	iterations	6	5	5	10
4 observations		multiple of Δ			
t_o		1	4	8	12
0	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	6.3e+1	2.9e+1	2.0e+2	4.6e+3
	iterations	4	4	4	4
4	δx_0	-0.1000	-0.1000	-0.1000	-0.0868
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1023
	κ	4.6e+2	8.1e+2	1.1e+4	2.4e+5
	iterations	4	4	5	10
8	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	2.9e+3	5.3e+3	7.4e+4	1.8e+6
	iterations	7	7	6	6
6 observations		multiple of Δ			
t_o		1	4	8	12
0	δx_0	-0.1000	-0.1000	-0.1000	-0.1000
	$\delta\alpha$	0.1000	0.1000	0.1000	0.1000
	κ	3.9e+1	2.9e+1	2.0e+2	4.6e+3
	iterations	3	4	4	4
4	δx_0	-0.1000	-0.1000	-0.1000	NaN
	$\delta\alpha$	0.1000	0.1000	0.1000	NaN
	κ	3.5e+2	8.1e+2	1.1e+4	1.0e+12
	iterations	4	4	5	10
8	δx_0	-0.1000	-0.1000	-0.1000	NaN
	$\delta\alpha$	0.1000	0.1000	0.1000	NaN
	κ	2.2e+3	5.3e+3	7.4e+4	7.2e+12
	iterations	7	7	6	10

Table 6: Estimates of β using the iterated second-order method. κ denotes the condition number of the matrix $\overline{H}_N^{\mathsf{T}}\overline{H}_N$, where N is the number of observations used in the estimation. The number of iterations is that required to reach a threshold of 10^-6 (at most 10)

| True Parameter Values: $x_o = 0.5$, $\alpha = 1.0$ | | | | | |
| Model Values | | | Corrections | | |
t_0	x_0	α	δx_0	$\delta \alpha$	Iterations
0	0.3	0.8	0.2000	0.2000	5
	0.3	0.9	0.2000	0.1000	5
	0.3	1.1	0.2000	-0.1000	5
	0.3	1.2	0.2000	-0.2000	5
	0.4	0.8	0.1000	0.2000	5
	0.4	1.1	0.1000	-0.1000	4
	0.4	0.9	0.1000	0.1000	4
	0.4	1.2	0.1000	-0.2000	4
	0.6	0.8	-0.1000	0.2000	4
	0.6	0.9	-0.1000	0.1000	4
	0.6	1.1	-0.1000	-0.1000	5
	0.6	1.2	-0.1000	-0.2000	5
	0.7	0.8	-0.2000	0.2000	4
	0.7	0.9	-0.2000	0.1000	5
	0.7	1.1	-0.2000	-0.1000	5
	0.7	1.2	-0.2000	-0.2000	5
4	0.3	0.8	0.2000	0.2000	7
	0.3	0.9	0.2000	0.1000	6
	0.3	1.1	0.2000	-0.1000	5
	0.3	1.2	0.2000	-0.2000	5
	0.4	1.1	0.1000	-0.1000	4
	0.4	1.2	0.1000	-0.2000	5
	0.4	0.8	0.1000	0.2000	6
	0.4	0.9	0.1000	0.1000	5
	0.6	0.8	-0.1000	0.2000	5
	0.6	0.9	-0.1000	0.1000	4
	0.6	1.1	-0.1000	-0.1000	6
	0.6	1.2	-0.1000	-0.2000	8
	0.7	0.8	-0.2000	0.2000	5
	0.7	0.9	-0.2000	0.1000	5
	0.7	1.1	-0.2000	-0.1000	7
	0.7	1.2	-0.2000	-0.2000	10
8	0.3	0.8	0.2000	0.2000	7
	0.3	1.1	0.2000	-0.1000	5
	0.3	0.9	0.2000	0.1000	7
	0.3	1.2	0.2000	-0.2000	7
	0.4	0.8	0.1000	0.2000	7
	0.4	0.9	0.1000	0.1000	6
	0.4	1.1	0.1000	-0.1000	5
	0.4	1.2	0.1000	-0.2000	11
	0.6	0.8	-0.1000	0.2000	6
	0.6	1.1	-0.1000	-0.1000	6
	0.6	1.2	NaN	NaN	100
	0.6	0.9	-0.1000	0.1000	5
	0.7	1.1	-0.2000	-0.1000	9
	0.7	1.2	NaN	NaN	100
	0.7	0.8	-0.2000	0.2000	6
	0.7	0.9	-0.2000	0.1000	5

Table 7: Iterated First-Order Corrections Using Different Model Parameters α and Initial Conditions x_0, in all cases using 4 observations spaced 0.5 seconds apart.

True Parameter Values: $x_o = 0.5$, $\alpha = 1.0$					
	Model Values		Corrections		
t_0	x_0	α	δx_0	$\delta \alpha$	Iterations
0	0.3	0.8	0.2000	0.2000	5
	0.3	0.9	0.2000	0.1000	5
	0.3	1.1	0.2000	-0.1000	5
	0.3	1.2	0.2000	-0.2000	5
	0.4	0.8	0.1000	0.2000	5
	0.4	0.9	0.1000	0.1000	5
	0.4	1.1	0.1000	-0.1000	4
	0.4	1.2	0.1000	-0.2000	4
	0.6	0.8	-0.1000	0.2000	4
	0.6	0.9	-0.1000	0.1000	4
	0.6	1.1	-0.1000	-0.1000	4
	0.6	1.2	-0.1000	-0.2000	4
	0.7	0.8	-0.2000	0.2000	4
	0.7	0.9	-0.2000	0.1000	4
	0.7	1.1	-0.2000	-0.1000	5
	0.7	1.2	-0.2000	-0.2000	5
4	0.3	0.8	NaN	NaN	101
	0.3	0.9	0.2000	0.1000	5
	0.3	1.1	0.2000	-0.1000	5
	0.3	1.2	0.2000	-0.2000	7
	0.4	0.8	0.1000	0.2000	7
	0.4	0.9	NaN	NaN	100
	0.4	1.1	0.1000	-0.1000	4
	0.4	1.2	0.1000	-0.2000	5
	0.6	0.8	-0.1000	0.2000	7
	0.6	0.9	-0.1000	0.1000	4
	0.6	1.1	-0.1000	-0.1000	5
	0.6	1.2	-0.1000	-0.2000	5
	0.7	0.8	-0.2000	0.2000	5
	0.7	0.9	-0.2000	0.1000	4
	0.7	1.1	-0.2000	-0.1000	5
	0.7	1.2	-0.2000	-0.2000	5
8	0.3	0.8	NaN	NaN	100
	0.3	1.1	0.2000	-0.1000	6
	0.3	1.2	0.2000	-0.2000	5
	0.3	0.9	0.2000	0.1000	16
	0.4	0.8	NaN	NaN	100
	0.4	0.9	NaN	NaN	100
	0.4	1.1	0.1000	-0.1000	5
	0.4	1.2	0.1000	-0.2000	4
	0.6	0.8	NaN	NaN	100
	0.6	0.9	-0.1000	0.1000	7
	0.6	1.1	-0.1000	-0.1000	5
	0.6	1.2	-0.1000	-0.2000	6
	0.7	0.8	-0.2000	0.2000	12
	0.7	0.9	-0.2000	0.1000	5
	0.7	1.1	-0.2000	-0.1000	5
	0.7	1.2	-0.2000	-0.2000	9

Table 8: Iterated Second-Order Corrections Using Different Model Parameters α and Initial Conditions x_0, in all cases using 4 observations spaced 0.5 seconds apart.